Systematic Error with Limited Statistics

- Some systematic errors on a top mass measurement are evaluated by using a difference of M_{top} estimates obtained from two statistically independent MC samples: one sample with nominal value of some parameter (*e.g.* amount of FSR) and the other with shifted value of that parameter. There seems to be a "convention" to use the larger value between the M_{top} shift and the statistical error on that shift as a systematic error due to that parameter.
- This "convention" is misguided because the final systematics consists of the sum in quadrature of several different systematic errors. In this case it is easy to show that always taking the shift already results in a conservative estimate of the systematics.
- The "convention" results in an inconsistent (in both common and statistical sense of the word) estimate of the systematics error. To see this, just use many sources of systematics of little relevance to the measured quantity. In this case, instead of combining the shifts (which can be close to 0), the "convention" combines the errors on the shifts (always finite) which results in a large systematic error no matter what the sources are.

Statistics of Using the Shifts

Suppose, the actual (unknown) shift of the M_{top} due to systematic source i is s_i, and it can be measured with precision σ_i and no bias. This presumes the use of two independent MC samples both of which are independent from the sample on which the analysis is tuned. Using the shifts results in an estimate of the systematic error squared which looks like

$$S^2 = \sum_i (s_i + e_i)^2$$

where the quantities e_i are distributed according to $N(0, \sigma_i)$ and are assumed to be uncorrelated. Then

$$\begin{aligned} E(S^2) &= \sum_i (s_i^2 + \sigma_i^2) \quad \text{Biased! Conservative!} \\ Var(S^2) &= E((S^2)^2) - (E(S^2))^2 = 2 \sum_i \sigma_i^2 (2 s_i^2 + \sigma_i^2) \end{aligned}$$

Assuming that all $\sigma_i \approx \sigma$,

$$\frac{\sqrt{Var(S^2)}}{E(S^2)} = \frac{\sqrt{2}\sigma \sqrt{\sum_i (2 s_i^2 + \sigma^2)}}{\sum_i (s_i^2 + \sigma^2)} < \frac{\sqrt{2}}{\sqrt{N}}$$

Improving Systematic Error Estimate

- One possible way to characterize the performance of a systematic error estimate is to find out how often the estimate S^2 is smaller than the real systematic error S_r^2.
- When the shifts are added in quadrature, the worst case performance happens when $N = 1$ and when s_1 is *large*. The method underestimates the systematic error 50% of the time. Note that the "convention" does not perform any better. For $N > 1$ the non-Gaussian shape of S^2 should be taken into account.
- Possible solution: choose how often we are willing to tolerate an underestimated systematic error. Let say, the underestimation probability we are willing to tolerate is α (perhaps, $\alpha = 0.05$ is reasonable). Then construct an improved estimate $S'^2 = S^2 - B(S^2) + \kappa(N, \alpha)\sqrt{Var(S^2)}$ which satisfies this performance requirement. Here, $B(S^2)$ is the bias of the S^2 estimate: $B(S^2) = \sum_i \sigma_i^2$, and $\kappa(N, \alpha)$ is a function that has to be tabulated from pseudoexperiments using the worst case scenario. $Var(S^2)$ has to be estimated in terms of the observed quantities. In terms of the observed shifts $s_i' = s_i + e_i$, a possible estimate is

$$S'^2 = \max\left(0, \sum_i s_i'^2 - \sum_i \sigma_i^2 + \kappa(N, \alpha)\sqrt{2\sum_i \sigma_i^2 (2s_i'^2 + \sigma_i^2)}\right)$$

JES Correlations

- Currently, we assume that the JES systematic errors are highly correlated for all jet p_T values. In general, this does not have to be the case. What we need is an estimate of how correlated these errors really are: the correlation function. Let's define

$$Var(p_{T1}, p_{T2}) = E[JES(p_{T1})JES(p_{T2})] - E[JES(p_{T1})]E[JES(p_{T2})]$$

Then

$$Corr(p_{T1}, p_{T2}) = \frac{Var(p_{T1}, p_{T2})}{[Var(p_{T1}, p_{T1})\, Var(p_{T2}, p_{T2})]^{\frac{1}{2}}}$$

To first order, this function represents our uncertainty about the linearity of the fully corrected jet response.

- For a finite number of p_T bins this becomes the correlation matrix.

- Ideally, the correlation function (or matrix) should be produced by the jet corrections group together with the JES systematic error. When the correlation function is available, we will be able to make much more precise statements about the systematic error on the top mass due to JES.

Differential Pulls

- The differential pull width is the calibration curve for statistical error estimates. It is built by splitting the pseudo experiments into a number of bins in the estimated error and calculating the pull width separately in each bin. An ideal differential pull width is 1 for all bins.
- If you think about it a little, the calibration curve for errors is as necessary as the calibration curve for the measured parameter. If you estimate the error from data then your error estimator is a random quantity and, therefore, it can exhibit an error-dependent bias. If you don't want to use a random error estimator then consistent use of frequentist statistics requires that you construct your error estimator from Monte Carlo rather than from data. This construction is not difficult but this is not what all top mass analyses are doing.
- A few years ago Jeremy Lys suggested that top mass analyses adjust their error estimates according to their differential pulls rather than standard (integrated) pulls. It is a good time to take heed of this suggestion and to make things right.